Chapter 5
Protein Purification and Characterization Techniques

SUMMARY

Section 5.1

- To begin the process of purification, proteins are released from cells using homogenization using a variety of physical techniques.
- Initial purifications steps are accomplished using differential centrifugation and salting out with ammonium sulfate.

Section 5.2

- Column chromatography refers to several common techniques for purification of proteins
- In gel filtration chromatography, proteins are separated by size
- In ion exchange chromatography molecules with a specific charge are selectively bound to a column, separated from proteins that don't bind, and then eluted.
- In affinity chromatography molecules are bound to the column via specific interactions for a bound ligand. Once non-binding proteins are removed, the protein of interest can be eluted.
- High Performance Liquid Chromatography (HPLC) exploits the same principles seen with other chromatographic techniques, but very high resolution columns that can be run under high pressures are used.
- Reverse Phase HPLC is a technique for the separation of nonpolar molecules. A solution of nonpolar compounds is put through a column that has a nonpolar liquid immobilized on an inert matrix. A more polar liquid serves as the mobile phase and is passed over the matrix. The solute molecules are eluted in proportion to their solubility in the more polar liquid.

Section 5.3

- Electrophoresis separates molecules on a gel medium by passing electrical current through the gel.
- Proteins are separated on the gel based on their size, shape, and charge.
- With SDS polyacrylamide gel electrophoresis, proteins separate based on molecular weight.

Section 5.4

- The amino acid sequence of a protein can be determined using a multi-step process.
- First, the protein is hydrolyzed into its constituent amino acids and the composition determined.
- The protein is also cleaved into smaller fragments and these fragments are then sequenced by the Edman Degradation.
- By using overlapping fragments and the sequences determined, the sequence of the original protein can be deduced.

- One of the most powerful techniques for determining protein structure is mass spectrometry (MS). A mass spectrometer exploits the difference in the mass to charge ratio (*m/z*) of ionized atoms or molecules in order to separate them from each other. The *m/z* ratio is such a characteristic property that it can be used to get structural and chemical information about the molecules and identify them.

Section 5.5

- Many techniques are used to separate and identify proteins.
- Enzyme-linked immunoabsorbent assays (ELISA) is a simple technique that involves screening proteins samples using a plate with wells in it that have an antibody specific for a protein being studied.
- A western blot is another technique using antibodies. In this case proteins are separated by gel electrophoresis and then the protein bands transferred to a thin membrane. The membrane is reacted with a specific antibody to identify where on the original gel the protein of interest was located.
- Protein chips take the same concept of an ELISA but allow for thousands of samples to be screened on one chip.

Section 5.6

- The proteome is the complete complement of proteins in a cell, and proteomics is the study of all of these proteins.
- Structural proteomics offers a detailed analysis of the structure of the proteins being produced.
- Expression proteomics analyzes the expression of proteins, and frequently considers their expression under different cellular conditions. It is a major contributor to our understanding of metabolism and disease.
- Interaction proteomics offers us the opportunity to look at how proteins interact with other molecules.
- A given cell will be producing thousands of proteins under different conditions. It takes many of the protein separation and identification techniques to determine the nature of the proteome.

LECTURE NOTES

This chapter will be most beneficial to students who are or will take a laboratory course in conjunction with the lecture course, but does contain interesting material regardless. In the course of the future chapters, the results obtained by researchers play a large role in the material presented, and this chapter will prepare the student to understand how researchers learn about biomolecules. Chapter 13 will continue with the techniques used in molecular biology, while this one focuses on those key techniques involved in protein analysis.

LECTURE OUTLINE

I. Protein purification
    A. Isolation of proteins from cells
        1. Cell fractionation
        2. Salting out
    B. Column chromatography
        1. Basic concepts – stationary vs. mobile phases
        2. Size-exclusion chromatography
        3. Affinity chromatography
        4. Ion-exchange chromatography
    C. Electrophoresis
        1. Agarose & polyacrylamide gels
        2. SDS-PAGE
        3. Isoelectric focusing
II. Primary structure determination
    A. Hydrolysis of proteins – HPLC analysis
    B. Sequence-specific cleavage
        1. trypsin
        2. chymotrypsin
        3. CNBr
    C. Edman degradation
III. Protein Identification Techniques
    A. ELISA
    B. Western Blot
    C. Protein Chips
IV. Proteomics

ANSWERS TO PROBLEMS

5.1 Extracting Pure Proteins from Cells

1. Using a blender, a Potter-Elvejhem homogenizer, or a sonicator.
2. If you needed to maintain the structural integrity of the subcellular organelles, a Potter-Elvejhem homogenizer would be better because it is more gentle. The tissue, such as liver, must be soft enough to use with this device.
3. Salting out is a process whereby a highly ionic salt is used to reduce the solubility of a protein until it comes out of solution and can be centrifuged. The salt forms ion–dipole bonds with the water in the solution, which leaves less water available to hydrate the protein. Nonpolar side chains begin to interact between protein molecules, and they become insoluble.
4. Their amino acid content and arrangements make some proteins more soluble than others. A protein with more highly polar amino acids on the surface is more soluble than one with more hydrophobic ones on the surface.
5. First homogenize the liver cells using a Potter-Elvejhem homogenizer. Then spin the homogenate at 500 × *g* to sediment the unbroken cells and nuclei. Centrifuge the supernatant at 15,000 × *g* and collect the pellet, which contains the mitochondria.
6. No, peroxisomes and mitochondria have overlapping sedimentation characteristics. Other techniques, such as sucrose-gradient centrifugation, would have to be used to separate the two organelles.
7. If the protein were cytosolic, once the cells were broken open, you could centrifuge at 100,000 × *g*, and all the organelles would be in the pellet. Your enzyme would be in the supernatant, along with all the other cytosolic ones.
8. Isolate the mitochondria via differential or sucrose-gradient centrifugation. Use another homogenization technique, combined with a strong detergent, to release the enzyme from the membrane.
9. Tables exist to tell you how many grams of ammonium sulfate [$(NH_4)_2SO_4$] to add to get a certain percent saturation. A good plan would be to take the homogenate and add enough ammonium sulfate to yield a 20% saturated solution. Let the sample sit for 15 minutes on ice and then centrifuge. Separate the supernatant from the precipitate. Assay both for the protein you are working with. Add more ammonium sulfate to the supernatant to arrive at a 40% saturated solution and repeat the process. In this way, you will find out what the percent saturation in ammonium sulfate needs to be to precipitate the protein.
10. Reasonably harsh homogenization would be able to liberate the soluble protein X from the peroxisomes, which are fragile. Centrifugation at 15,000 × *g* would sediment the mitochondria (broken or intact). The supernatant would then have protein X but no protein Y. Freeze/thaw techniques and sonication would accomplish the same thing, or the mitochondria and the peroxisomes could be separated initially by sucrose-gradient centrifugation.

5.2 Column Chromatography

11.
(a) Size.
(b) Specific ligand-binding ability.
(c) Net charge.
(d) Polarity

12. The largest proteins elute first; the smallest elute last. Larger proteins are excluded from the interior of the gel bead so they have less available column space to travel. Essentially, they travel a shorter distance and elute first.
13. A compound can be eluted by raising the salt concentration or by adding a mobile ligand that has a higher affinity for the bound protein than the stationary resin ligand does. Salt is cheaper but less specific. A specific ligand may be more specific, but it is likely to be expensive.
14. A compound can be eluted by raising the salt concentration or by changing the pH. Salt is cheap, but it might not be as specific for a particular protein. Changing the pH may be more specific for a tight pI range, but extremes of pH may also denature the protein.
15. Raising the salt concentration is relatively safe. Most proteins will elute this way, and, if the protein is an enzyme, it will still be active. If necessary, the salt can be removed later via dialysis. Changing the pH enough to remove the charge can cause the proteins to become denatured. Many proteins are not soluble at the isoelectric points.
16. The basis of most resins is agarose, cellulose, dextran, or polyacrylamide.
17. See Figure 5.7.
18. Within the fractionation range of a gel-filtration column, molecules elute with a linear relationship of log MW versus their elution volumes. A series of standards can be run to standardize the column, and then an unknown can be determined by measuring its elution volume and comparing it to a standard curve.
19. Both proteins would elute in the void volume together and would not be separated.
20. Yes, the β-amylase would come out in the void volume, but the bovine serum albumin would be included in the column bead and would elute more slowly.
21. In most chromatography systems, the ligands and solvents are polar. In reverse phase HPLC, a solution of nonpolar compounds is put through a column that has a nonpolar liquid immobilized on an inert matrix. A more polar liquid serves as the mobile phase and is passed over the matrix. The solute molecules are eluted in proportion to their solubility in the more polar liquid.
22. Ion Exchange Chromatography is a specific type of separation based on net charge of the molecules being separated. The term HPLC refers to chromatography procedures carried out under high pressure, but the basis of the separation could be ion exchange, gel filtration, reverse phase, or affinity chromatography.

23. Set up an anion-exchange column, such as Q-Sepharose (quaternary amine). Run the column at pH 8.5, a pH at which the protein X has a net negative charge. Put a homogenate containing protein X on the column and wash with the starting buffer. Protein X will bind to the column. Then elute by running a salt gradient.
24. Use a cation-exchange column, such as CM-Sepharose, and run it at pH 6. Protein X will have a positive charge and will stick to the column.
25. With a quaternary amine, the column resin always has a net positive charge, and you don't have to worry about the pH of your buffer altering the form of the column. With a tertiary amine, there is a dissociable hydrogen, and the resin may be positive or neutrally charged, depending on the buffer pH.
26. The easiest way would be to use a sucrose gradient to separate the mitochondria from the peroxisomes first. Then break open the mitochondria via harsh homogenization or sonication, and then centrifuge the mitochondria. The pellet would contain protein B, while the supernatant would contain protein A. Contaminants could still exist, but they could be cleaned away by running gel filtration, on Sephadex G-75 (which would separate enzyme C from enzymes A and B), and then by running ion-exchange chromatography on Q-Sepharose at pH 7.5. Enzyme B would be neutral and would elute, while enzyme A would stick to the column.
27. Glutamic acid will be eluted first because the column pH is close to its pI. Leucine and lysine will be positively charged and will stick to the column. To elute leucine, raise the pH to around 6. To elute lysine, raise the pH to around 11.
28. A nonpolar mobile solvent will move the nonpolar amino acids fastest, so phenylalanine will be the first to elute, followed by glycine and then glutamic acid.
29. The nonpolar amino acids will stick the most to the stationary phase, so glutamic acid will move the fastest, followed by glycine and then phenylalanine.
30. A protein solution from an ammonium sulfate preparation is passed over a gel-filtration column where the proteins of interest will elute in the void volume. The salt, being very small, will move through the column slowly. In this way, the proteins will leave the salt behind and exit the column without it.

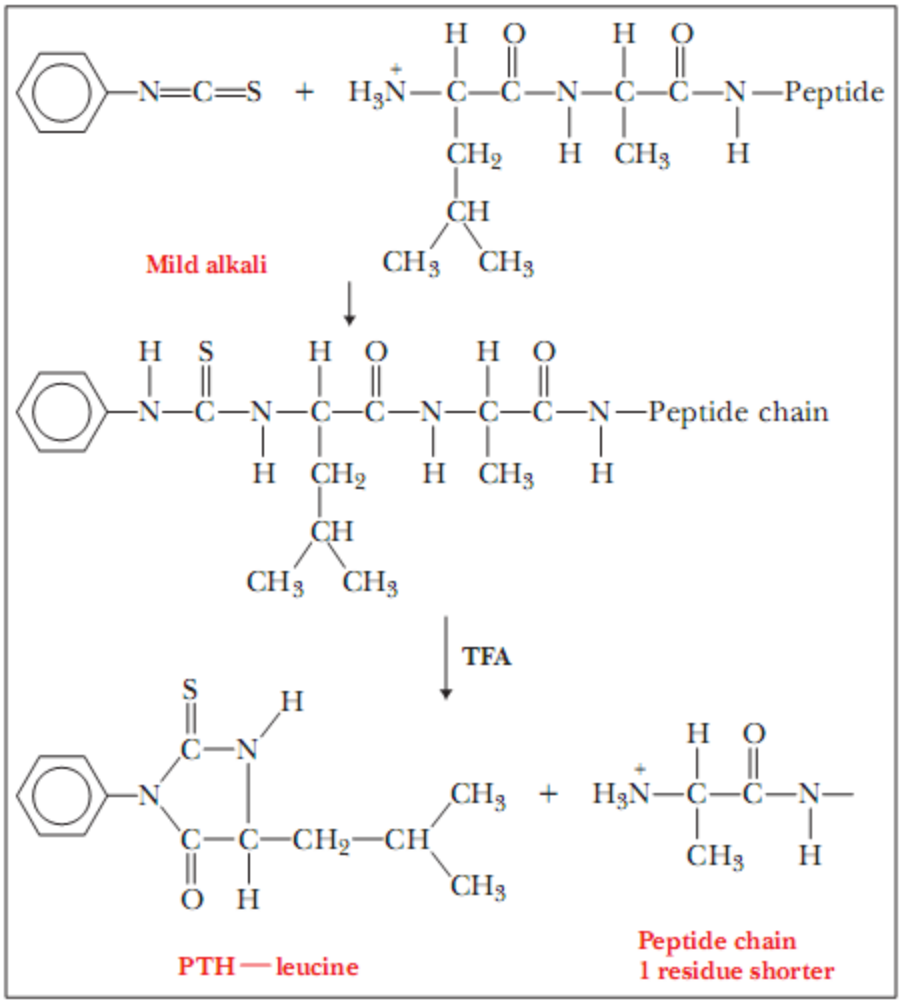
5.3 Electrophoresis

31. Size, shape, and charge.
32. Agarose and polyacrylamide.
33. Polyacrylamide.
34. DNA is the molecule most often separated on agarose electrophoresis, although proteins can also be separated.
35. Those with the highest charge/mass ratio would move the fastest. There are three variables to consider, and most electrophoreses are done in a way to eliminate two of the variables so that the separation is by size or by charge, but not by both.
36. Sodium dodecyl-sulfate polyacrylamide gel-electrophoresis. With SDS–PAGE, the charge and shape differences of proteins are eliminated so that the only parameter determining the migration is the size of the protein.

37. SDS binds to the protein in a constant ratio of 1.4 g SDS per gram of protein. It coats the protein with negative charges and puts it into a random coil shape. Thus, charge and shape are eliminated.
38. In a polyacrylamide gel used for gel-filtration chromatography, the larger proteins can travel around the beads, thereby having a shorter path to travel and therefore eluting faster. With electrophoresis, the proteins are forced to go through the matrix, so the larger ones travel more slowly because there is more friction.
39. The MW is 37,000.

5.4 Determining the Primary Structure of a Protein

40. The Edman degradation will give the identity of the N-terminal amino acid in its first cycle, so doing a separate experiment is not necessary.
41. It might tell you if the protein were pure or if there were subunits.
42.



43. The amount of Edman reagent must exactly match the amount of N-termini in the first reaction. If there is too little Edman reagent, some of the N-termini will not react. If there is too much, some of the second amino acid will react. In either case, there will be a small amount of contaminating phenylthiohydantoin (PTH) derivatives. This error grows with the number of cycles run until the point that two amino acids are released in equal amounts, and you cannot tell which one was supposed to be the correct one.
44. In the first cycle, the first and second amino acids from the N-terminal end would be reacted and released as PTH derivatives. You would get a double signal and not know which one was the true N-terminus.
45. Val—Leu—Gly—Met—Ser—Arg—Asn—Thr—Trp—Met—Ile—Lys—Gly—Tyr—Met—Gln—Phe
46. Met—Val—Ser—Thr—Lys—Leu—Phe—Asn—Glu—Ser—Arg—Val—Ile—Trp—Thr—Leu—Met—Ile
47. It is possible that your protein is not pure and needs additional purification steps to arrive at a single polypeptide. It is also possible that the protein has subunits, so multiple polypeptide chains could be yielding the contradictory results.
48. There are two fragments that have C-termini that are not lysine or arginine, which is what trypsin is specific for. Normally there would be only one fragment ending with an amino acid that was not Arg or Lys, and we would immediately know that it was the C-terminus. Histidine is a basic amino acid, although it is usually neutral and therefore does not react with trypsin. It is possible that, in the pH environment of the reaction, the histidine was positively charged and was recognized by trypsin.
49. It would tell you a relative concentration of the various amino acids. This is important because it would help you plan your sequencing experiment better. For example, if you had a protein whose composition showed no aromatic amino acids, it would be a waste of time to use a chymotrypsin digestion.
50. Cyanogen bromide would be useless, because there is no methionine. Trypsin would be little better, because the protein has 35% basic residues. Trypsin would shred the protein into more than 30 pieces, which would be very hard to analyze.
51. Chymotrypsin would be a good choice. There are more than four residues of aromatic amino acids. The protein, containing 100 amino acids, would be cut four times, possibly yielding nice fragments roughly 20–30 amino acids long, which can be sequenced effectively by the Edman degradation.
52. It would work best if the aromatic residues were spread out in the protein. In that way, fragments in the proper size range would be generated. If all four of the aromatic residues were in the first 10 amino acids, there would be one long fragment that could not be sequenced.
53. Electrospray Ionization (ESI-MS) and Matrix-Assisted Laser Desorption Ionization- Time of Flight (MALDI-TOF MS).
54. MALDI-TOF MS is very sensitive and very accurate. Attomole (10^{-18}) quantities of a molecule can be detected.

5.5 Protein Identification Techniques

55. ELISA is based on antibody-protein interactions. Specific antibodies, called primary antibodies, are put into microtiter plates to localize target proteins. A secondary antibody carrying some type of tag to render it visible is also added. If the target protein is there, then the combination of the protein-primary antibody-secondary antibody will be visible.
56. A primary antibody is specific for a target protein that a researcher is looking for. A secondary antibody will react with the primary antibody. The secondary antibody carries the tag that makes the complex visible.
57. The protein-antibody complexes can be seen based on the nature of the tag carried by the secondary antibody. This tag may be an enzyme that produces a visible color when provided its substrates, a fluorescent marker, or a radioactive compound.
58. The first step in a western blot is the separation of proteins via electrophoresis. The next step takes the gel from the electrophoresis and transfers the proteins from the gel onto a thin membrane of nitrocellulose or other absorbing compound. Once the proteins are transferred, they are incubated with the primary antibody. Then they are incubated with the secondary antibody. Lastly, the bands are made visible by reacting with the substrates for the secondary antibody enzyme tag, or they are visualized with a fluorometer or x-ray paper.
59. Western blot got its name as a humorous derivation from the original blotting technique called Southern blotting. The original blotting technique was for DNA and was developed by a researcher named Southern, so they called the technique the Southern blot. The next type of molecule to be blotted was RNA, so to distinguish from a Southern blot, it was called a northern blot. After that, the technique for blotting protein was developed and called a western blot.
60. The advantages of an ELISA would be ease of use, low cost, and ready availability to any researcher. The disadvantages are that, compared to a microarray, relatively few proteins can be tested at one time. A microarray can test thousands of proteins in a single experiment, so it is much more powerful. However, it is also much more expensive and requires specialized equipment not readily available.
61. Proteins are transferred to nitrocellulose because all of the protein ends up layered on top of a very thin membrane. This means that small volumes of the antibody solutions can be used to bind to the proteins. These antibodies are very expensive, so the less used, the better. Also, since the original proteins were imbedded in a gel, if the gel were reacted directly with the antibodies, the antibodies would not have easy access to the gel, as they would not percolate easily through the pores of the gel without the benefit of an electric current to push them.
62. There are thousands of primary antibodies that can be purchased commercially. Other research requires a new primary antibody to be created. The process of attaching an enzyme, a fluorescent marker, or a radioactive compound is a long and difficult task. If every primary antibody had to be tagged, it would be a daunting exercise for the company trying to do it. Instead, a company can

specialize in tagging a secondary antibody directed against goat, rabbit, or mouse antibodies. These secondary antibodies can then be used for any experiment where the primary antibody came from one of these animals. Also, there is a signal multiplication effect of using a primary antibody since several secondary antibodies will attach to a single molecule of primary antibody.

5.6 Proteomics

63. Proteomics is the systematic analysis of an organism's complete complement of proteins, or its proteome. Just as we learned the basic dogma of molecular biology (DNA →RNA → protein), the technology now available has allowed scientists to describe all the DNA of an organism as its genome, all of the RNA as its transcriptome, and all of the proteins produced as its proteome. To understand the flux of proteins in a cell is to understand its metabolism.
64. The bait protein is constructed to have a particular affinity tag. The bait protein interacts with cell proteins of interest and then binds to an affinity column via the tag. In this way, the cell proteins of interest can be found and isolated.
65. There are many assumptions behind the experiment described in the Biochemical Connections on page135. One must assume that the nature of the tag has not changed the binding of the protein. For example, if adding the tag makes a protein more likely or less likely to bind to it, then the conclusions about cellular protein binding may be incorrect. For example, one might conclude that two proteins bind together to serve their metabolic function, but this binding could be an artifact of the experimental conditions. One must also assume that tagging the proteins has not changed the affinity between the tag and the affinity column.